Peer-reviewed version available at *Remote Sens.* **2020**, *12*, 118; doi:10.3390/rs12010118

*Article*

# Assessing OpenStreetMap Completeness for Management of Natural Disaster by Means of Remote Sensing: A Case Study of Three Small Island States

Ran Goldblatt [1,*], Nicholas Jones [2] , Jenny Mannix [3]

[1] New Light Technologies Inc.; ran.goldblatt@nltgis.com
[2] World Bank Group; njones@worldbank.org
[2] New Light Technologies Inc.; jennifer.mannix@nltgis.com
* Correspondence: ran.goldblatt@nltgis.com; Tel.: +1-202-630-0362

**Abstract:** Over the last few decades, many countries, especially Caribbean island ones, have been challenged by the devastating consequences of natural disasters, which pose a significant threat to human health and safety. Timely information related to the distribution of vulnerable population and critical infrastructure are key for an effective disaster relief. OpenStreetMap (OSM) has repeatedly been shown to be highly suitable for disaster mapping and management. However, large portions of the world, including countries exposed to natural disasters, remain unmapped. In this study, we propose a methodology that relies on remotely sensed measurements (e.g. VIIRS, Sentinel-2 and Sentinel-1) and derived classification schemes (e.g. forest and built-up land cover) to predict the completeness of OSM building footprints in three small island states (Haiti, Dominica and St. Lucia). We find that the combinatorial effects of these predictors explain up to 94% of the variation of the completeness of OSM building footprints. Our study extends the existing literature by demonstrating how remotely sensed measurements could be leveraged to evaluate the completeness of OSM database, especially in countries at high risk of natural disasters. Identifying areas that lack coverage of OSM features could help prioritize mapping efforts, especially in areas vulnerable to natural hazards and where current data gaps pose an obstacle to timely and evidence-based disaster risk management actions.

**Keywords:** OpenStreetMap; Disaster Management; OSM Coverage, Remote Sensing.

## 1. Introduction

Over the last few decades, many countries have been challenged by the devastating consequences of natural disasters which pose a significant threat to human health and safety and impact vulnerable communities and critical infrastructure globally. Every year, natural disasters impact close to 160 million people worldwide (WHO, 2019), causing destruction of the physical, biological and social environments, impacting food security, and causing global losses that amount to over 100 billion dollars (FAO, 2019). The frequency of natural disasters has been steadily increasing since 1940 [3] and over the next century climate change will likely amplify the number and severity of such disasters [4]. By 2030, up to 325 million extremely poor people will live in the 49 most hazard prone countries, the majority of them in South Asia and sub-Saharan Africa [5].

While the impacts of natural disasters are worldwide, some countries have been more vulnerable to different types of disasters than others [6]. For example, in 2017 Puerto Rico, Sri Lanka and Dominica were at the top of the list of the most affected countries to natural disasters such as significant precipitation, floods and landslides. Of the ten most affected countries and territories to natural disasters since 1998, eight were developing countries in the low income or lower-middle income country group; one was an upper-middle income country (Dominica); and one an advanced economy (Puerto Rico) [7]. Caribbean island countries are especially exposed to a wide range of

natural disasters, such as tropical cyclones, tsunamis, flooding, volcanic eruptions and earthquakes [8] and small island developing states - which are frequently characterized by coastal communities, geographic isolation, and limited technical capacity - are among the most vulnerable countries to natural disasters and climate change [9].

Recognizing these trends, there is an increasing need for efficient and well-planned disaster management and disaster relief operations. The term disaster risk management refers to the full lifecycle of actions aiming to prevent, prepare for, respond to, and recover from disasters. Generally, disaster risk management consists of four main phases: (1) *Mitigation*, i.e. activities that reduce the likelihood and expected adverse impacts of a natural disaster event; (2) *Preparedness*, i.e. plans or preparations to strengthen emergency response capabilities; (3) *Response*, i.e. actions taken to save lives and prevent property damage in an emergency situation; and (4) *Recovery*, i.e. interventions aimed at returning communities and infrastructure to a proper level of functionality following a disaster.

Timely geospatial information related to the distribution of vulnerable population and the location, availability and functionality of critical infrastructure (e.g. hospitals, shelters, water and sanitation facilities, roads and public transportation) is key for an effective disaster relief not only during the response phase, but also to support post-event recovery efforts and to guide an effective planning in preparation for the next disaster.

Until recently, governmental agencies and the commercial sector were the primary sources for geospatial data for disaster management. In the past decade, however, the public has been increasingly recognized as a valuable source for geospatial information for disaster management [10]. Recent developments in web mapping technologies have led to disaster management operations that are more dynamic, transparent, and decentralized, increasingly with contribution by individuals and organizations from both inside and outside the impacted area [11] including by means of geospatial information that is contributed by volunteers.

The term Volunteered Geographic Information (VGI) refers to geographic information collected by individuals also known as Volunteer and Technical Communities (VTC) or digital humanitarians [12], often on a voluntary basis [10]. The trend of VGIs was recognized by *Time* magazine in 2006, when "You" were chosen as the magazine's *Person of the Year*. VGIs allow to rapidly collect accurate information before, during and following a disaster, making this information open and freely accessible [13] and filling deficiencies of traditional mapping technologies and sources of data [14] [15].

*OpenStreetMap (OSM) for Disaster Management*

Created in 2004 by Steve Coast, OpenStreetMap (OSM) is a collaborative project aiming to "create a free editable map of the world". During its first year, most mapping efforts focused on road and transportation networks. Today, a variety of geographical features, including buildings and their functionality, land use and public transportation information are constantly added to OSM [16]. Such data allows local governments and communities to better perform risk assessment and emergency planning [17] [18] [19] and is routinely utilized for various disaster risk management applications [20] [21]. As of today, there are more than 5.5 million OSM users and one million contributors who generate more than 3 million changes every day, as well as specialized groups such as Humanitarian OpenStreetMap Team (HOT-OSM) that conduct activities aimed at enriching OSM data to support emergency relief operations[i].

The need for comprehensive up-to-date geospatial information is especially evident during and immediately after major disaster events. In the days following a magnitude-7.0 earthquake in Haiti (January 12, 2010), there was lack of this fundamental data. In response to the event, numerous organizations have released high-resolution satellite imagery under open licensing schemes, catalyzing worldwide efforts to map Haiti in order to support the recovery operations. Over 450 OSM volunteers from 29 countries relied on this imagery to digitize roads, buildings, and other features in the affected areas, creating the most detailed map of Haiti in existence in just a few weeks [22]. Haiti has been subject to a wide range of severe disaster events. For example, on October 4, 2016, Hurricane

Matthew struck southwestern Haiti, leaving 900 people killed and 28,000 homes damaged. In response to the event, significant updates to OSM database occurred especially for four days after the hurricane hit the island [23].

*OpenStreetMap (OSM) Data Completeness and Accuracy*

Although OSM road network data is estimated to exceed 80% completeness in relation to the world's roads and streets [24], in general, the coverage and completeness of OSM features (including building footprints) varies significantly - not only between countries, but also within countries. For example, completeness of coverage of remote areas is often lower than of highly populated urban areas, and the coverage of developed countries tends to be lower than of developing countries [25] [26] [27] [28] [29]. These differences are due, in part, to societal factors, such as population distribution and population density, distance to major cities and the location of contributing users [26] [30] [31] [32] [33] [34].

With the increased utilization of VGIs for disaster preparedness and response, various methodologies have been proposed to assess the quality and the accuracy of VGIs [15], for example, in terms of data completeness, logical consistency, positional, thematic, semantic and spatial accuracy, temporal quality and usability [25] [35] [36] [37] [38] [39] [36].

In this study, we propose a methodology that utilizes remotely sensed observations to predict the coverage of OSM features and to identify gaps in the completeness of OSM building footprints. In the past, expensive satellite imagery and limited computational power only allowed analysis of small geographical contexts, for example, counting building footprints in a small neighborhood or evaluating the volume of live vegetation in a single agriculture field. This model is being replaced thanks to the availability of publicly available and free satellite data that capture every location on earth every few days and in a spatial resolution of up to a few meters, allowing to understand not only the physical characteristics of Earth, but also fundamental socio-economic patterns and processes. The availability of daytime (e.g. Sentinel, Landsat) and nighttime (e.g. DMSP-OLS, VIIRS) satellite imagery, together with advancements in the capabilities of cloud-based computational platforms, now allows for analysis of Land Cover and Land Use (LCLU) characteristics of Earth across a greater geographic and temporal scale. Different types of LCLU, including built-up areas and built-up structures, hold unique reflectance characteristics that can be differentiated by means of remote sensing. Our objective here is to rely on remotely sensed measurements to predict the coverage of OSM building footprints. Previous studies have utilized OSM data for different remote sensing applications, for example, for classification of urban areas [40] or for semantic labeling of aerial and satellite images [41]. Despite significant progress in the field of machine learning and the increasing availability of satellite imagery, there is still a scarcity of studies aiming to utilize remotely sensed observations to predict the completeness of OSM building footprints. Identifying areas that lack coverage of OSM features could help prioritize mapping efforts, especially in areas vulnerable to natural hazards and where current data gaps pose an obstacle to timely and evidence-based disaster risk management actions. We demonstrate our methodology in the case study of Haiti, one of the most vulnerable countries to natural disasters, and where there are still significant gaps in OSM completeness. We validate the methodology and our approach in the case of two additional small island states: Dominica and St. Lucia.

The remainder of this article is organized as follows. In Section 2, we discuss the methodology, the study area and the data we use to predict the coverage of OSM building footprints. In Section 3 we present and evaluate the results and in Sub-Section 3.3 we illustrate the applicability of our approach in the case of two small island states: Dominica and St. Lucia. In Section 4, we offer a concluding discussion.

## 2. Materials and Methods

### *2.1. Haiti*

Located on the western side of Hispaniola Island, Haiti (27,750 square kilometers in size, with a population of approximately 11.5 million) is the poorest country in the Western Hemisphere, with a Gross Domestic Product (GDP) per capita of US$ 870 [42]. Haiti is highly vulnerable to natural disasters; more than 96% of its population is exposed to different types of natural hazards, particularly hurricane, coastal and riverine flood, and earthquake [42]. More than half of the population lives in cities and towns, a major shift from the 1950s, when approximately 90% of Haitians lived in the countryside [43]. Almost all of Haiti's 30 major watersheds experience significant flood events, due to intense seasonal rainfall, storm surge in the coastal zones, deforestation and erosion, and sediment-laden river channels [44]. Furthermore, large portions of Haiti`s population (e.g. in the capital Port-au-Prince) live in shanty towns built upon steep and exposed hillsides [45]. In 2018 alone, some 2.8 million people were considered to be in need of humanitarian assistance valued at US$ 252.2 million [46].

### *2.2. Data*

#### 2.2.1 OSM data

As of today, approximately 930,000 buildings (totaling an area of 64.6 square Kilometers) and 20,948 km of roads have been added to OSM database over Haiti. To assess the current coverage of OSM building footprints in the country, we download the most recent dataset from Geofabrik (https://www.geofabrik.de/data/download.html) in Shapefile format (data downloaded in July 2019). For Dominica and St. Lucia, we download the most recent OSM data using overpass turbo (https://overpass-turbo.eu/), a web-based data filtering tool for OpenStreetMap (data downloaded in July 2019). We select only OSM features that are labeled "building" (38,619 and 29,412 polygons in Dominica and St. Lucia, respectively).

#### 2.2.2. Predictors

We create an artificial tessellated grid of cells that spans the entire country, each cell 0.25 Square Km in size, (a total of 136,747 cells).

We predict the area of OSM building footprints in a cell based on several remotely sensed measures and geospatial features related to the street network:

Nighttime lights (VIIRS): The Visible Infrared Imaging Radiometer Suite (VIIRS) is one of the key instruments onboard the Suomi National Polar-Orbiting Partnership (Suomi NPP) spacecraft (launched in 2011). VIIRS instrument collects visible and infrared imagery and global observations of land, atmosphere, cryosphere and oceans. This instrument has significant improvements over the capabilities of the former DMSP-OLS [47], notably its availability on a daily basis and higher spatial resolution (up to 500m at the equator). First, we record for each pixel the maximum value of all overlapping pixels (in the same location) in a stack of 10 monthly composites (Jan – Oct) of 2019. Then, for each cell we calculate a Sum of Light (SOL) measure (the sum of the digital number values of all overlapping pixels in each cell).

Sentinel-2 derived spectral indices: The Copernicus Sentinel-2 mission comprises a constellation of two polar-orbiting satellites that collect multi-spectral data in 13 spectral bands, with four bands at a spatial resolution of 10m and 6 bands at a spatial resolution of 20 m. The revising period of Sentinel-2 is 5 days at the equator. We calculate four remotely sensed measures sensitive to vegetation and built-up land cover: Normalized Difference Vegetation Index (*NDVI*) [48], Soil Adjusted Vegetation Index (*SAVI*) [49], Normalized Difference Built-up Index (*NDBI*) [50] and Urban Index (*UI*) [51] (Table 1).

Sentinel-1 SAR: Sentinel-1 mission comprises a constellation of two polar-orbiting satellites, operating day and night performing C-band synthetic aperture radar imaging, enabling them to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

acquire imagery regardless of the weather. Sentinel 1 revisit period is every 6 days, with a spatial resolution down to 5m. Similarly to [52], we capture the texture of the surface by utilizing Sentinel-1's C-band (single co-polarization vertical transmit and vertical receive (VV) acquisition mode). From each scene we remove speckle noise and perform radiometric calibration and terrain correction.

Slope: To capture the topography of the surface, we use the Global SRTM mTPI dataset, where a local gradient is calculated for each pixel based on Global SRTM elevation data.

Forest cover: We estimate the extent of forest cover in 2018 based on the Hansen Global Forest Change v1.6 (2000-2018). First, we define a pixel as "forest" in the year 2000 if more than 20% of it was covered in 2000 with forest. We record pixels that experienced a major event of forest cover loss between 2000 and 2018 and estimate the total area of forest cover in 2018 in each cell.

Urban footprints: We rely on two remotely sensed derived products signifying urban and rural settlements that were produced by the Earth Observation Center at DLR: The Global Urban Footprint (GUF) (in a spatial resolution of ~12m) and the World Settlement Footprint (WSF) (in a spatial resolution of ~10m) [53] [54] [55].

OSM transportation network features: We calculate the total length of OSM roads in a cell and the total number of junctions in a cell as additional potential predictors of OSM building footprints.

**Table 1.** The predictors used to predict per-cell area of OSM building footprints.

| Predictor | Source | Per-cell statistics |
|---|---|---|
| Nighttime lights | VIIRS | Sum of Light (SOL): The sum of $DN_{max}$ value of all pixels in cell, where $DN_{max_i}$ is the maximum digital number (DN) value of pixel in location $i$ over 12 monthly composites in 2019. |
| NDVI (NIR-RED)/ (NIR+RED) | Sentinel-2 | The sum NDVI value of all pixels in a cell |
| SAVI (NIR-RED) / (NIR+RED+L) * (1+L) | Sentinel-2 | The sum SAVI value of all pixels in a cell |
| NDBI (MIR-NIR)/ (MIR+NIR) | Sentinel-2 | The sum NDBI value of all pixels in a cell |
| UI (SWIR2-NIR)/ (SWIR2+ NIR) | Sentinel-2 | The sum UI value of all pixels in a cell |
| deforestation | Hansen Global Forest Change v1.6 (2000-2018) | Total forest cover in a cell (2018) |
| Built-up area | GUF | Total built up area in a cell |
| Built-up area | WSF | Total built up area in a cell |
| Topography (slope) | SRTM | Average slope per cell |
| Surface texture | Sentinel-1 | Average texture per cell |
| Roads | OSM | Total length of roads in a cell |
| Roads junctions | OSM | Number of junctions in a cell |

## 3. Results

### *3.1. Explanatory variables*

An examination of the 136,747 cells spanning Haiti shows that only 25.1% of the cells have at least one mapped building, and only 512 of them have more than 10% of their area covered with building footprints (Figure 1 shows a histogram of the distribution of OSM building footprints per cell). On average, there are 27.5 buildings in a cell (Std= 83.4); 1,530 of the cells (i.e. only 1.1% of the cells spanning Haiti) incorporate more than 100 mapped buildings. In comparison, 8.15% and 6.84% of the cells incorporate built-up land cover according to WSF and GUF, respectively.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

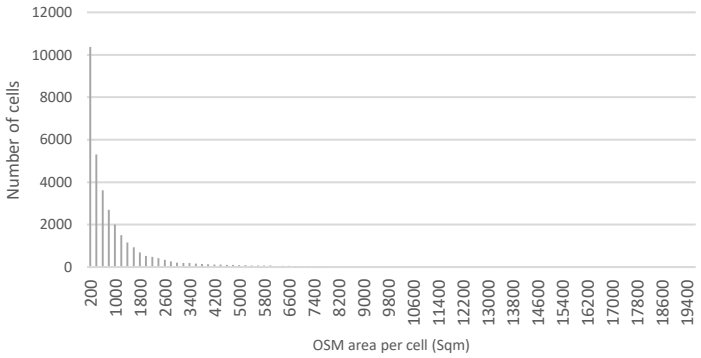



**Figure 1.** The distribution (histogram) of OSM building footprints area (Sqm) per cell

As discussed above, a visual examination of the completeness of OSM building footprints over Haiti suggests that large portions of the island remain unmapped (Figure 2a). Figure 2c shows, as an illustration, the coverage of OSM building footprints in the capital of Haiti, Port-au-Prince, and in the adjacent Carrefour commune. While buildings in many areas within these cities have been mapped, large portions are still not fully mapped. We observe that densely mapped zones of Port-au-Prince co-exist alongside zones that remain entirely unmapped (Figure 2b), a visual pattern that may result from episodic engagement of community mapping volunteers and the definition of mapping 'tasks' on a neighborhood scale through OSM editing tools. Moreover, significant parts in northern Haiti are not mapped (Figure 3), including, for example, the cities Gonaïves and Cap-Haitien.

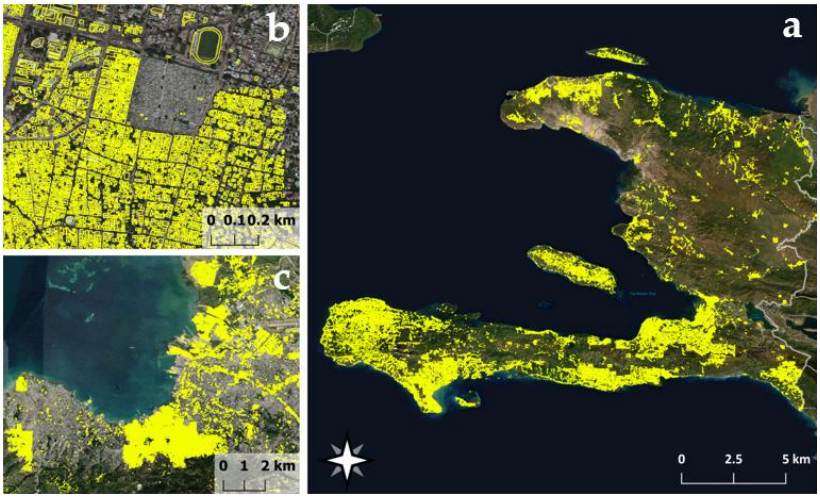
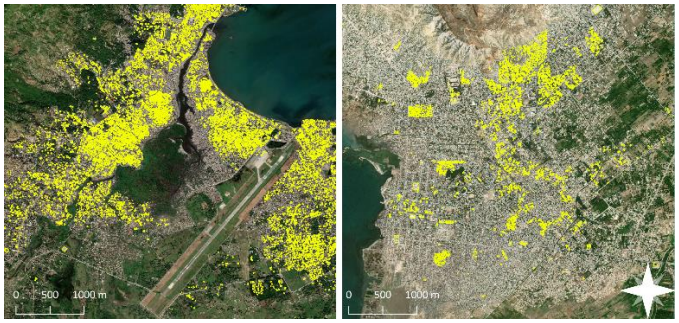



**Figure 2.** OSM building footprints coverage in Haiti (a), in the capital of Haiti, Port-au-Prince (b), and in the adjacent Carrefour commune (c)




(a) (b)

**Figure 3.** OSM building footprints coverage in (a) the city of Cap-Haitien and (b) Gonaïves in northern Haiti.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

A Pearson correlation test indicates a significant ($p<0.01$) correlation between the total area of OSM building footprints in a cell and several of the examined predictors. As expected, there is a positive and significant correlation between the area of OSM building footprints in a cell and the total area of built-up land cover according to WSF and GUF ($r=0.73$ and 0.71, respectively, $p<0.01$) as well as with nighttime lights (VIIRS SOL) ($r=0.63$, $p<0.01$). We find a significant ($p<0.01$) correlation between OSM building footprints area in a cell with the four Sentinel-2 spectral indices, indicated by a positive correlation with UI and NDBI ($r=0.59$ and $r=0.47$) and a negative correlation with both SAVI and NDVI ($r=-0.53$).

We visually assess the completeness of OSM building footprints in Haiti by overlaying the OSM building footprint dataset with the most recent high-resolution base map image (provided by ESRI). We identify 835 cells where we assess that at least 75% of the buildings are mapped (we choose a threshold of 75% because we want to sample cells representing diverse regions and geographies in Haiti, many of them have not been fully mapped) (Figure 4 shows examples of cells where more than 75% of the structures are mapped). The correlation between the area of OSM building footprints in a cell and the examined predictors is higher compared to the previous experiment, where all the cells where considered (for example, $r=0.78$ and $r=0.65$ with WSF and VIIRS and $r= 0.61$ and $r=-0.55$ with UI and SAVI, respectively) (Table 2). The higher correlation is likely due to the "noise" caused in the first experiment due to the incomplete coverage of OSM building footprints within the entire Haiti dataset (i.e. cells that incorporate building footprints but that have not been mapped).

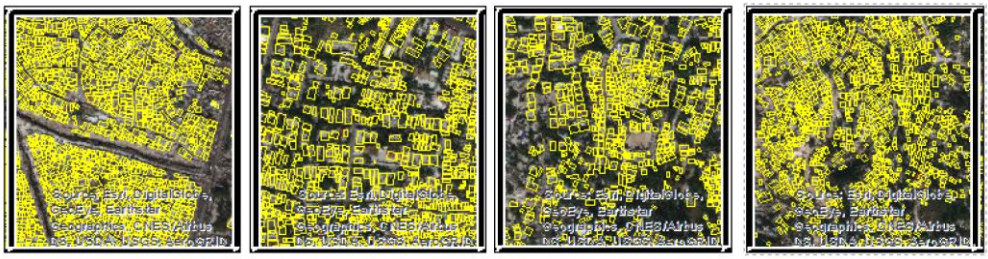


**Figure 4.** Examples of cells that have more than 75% of their area mapped with OSM building footprints.

**Table 2.** Pearson correlation test between the area of OSM building footprint in a cell and the evaluated predictors (only cells that were assessed as mapped, N=835).

| | VIIRS | GUF | WSF | NDVI | NDBI | SAVI |
|---|---|---|---|---|---|---|
| r | 0.654* | 0.76 | 0.78 | -0.551 | 0.486 | -0.551 |
| | UI | Forest Cover | SE1 | Slope | Road length | OSM junctions |
| r | 0.614 | -0.388 | 0.16 | -0.11 | 0.69 | 0.60 |

Note: *$p<0.01$

To evaluate the combinatorial effect of the predictors on the completeness of OSM building footprints, we perform an Ordinary Least Squares regression (OLS). OLS regression model describes the population of a random variable Y and assumes that the observations $Y_1$, $Y_2$, . .. , $Y_T$ constitute a random sample with a linear conditional mean made up of variables X, $X_2$, . .. ,$X_K$. The analysis shows that nine of the variables together explain up to 82% of the variation of OSM building footprint area in a cell ($R^2 = 0.82$, $F(12,822) = 323.20$, $p < 0.01$)(Table 3). We evaluate the contribution of four types (groups) of features to the model fit using a stepwise regression analysis: (1) only GUF and WSF; (2) with the addition of nighttime lights (VIIRS); (3) with the addition of further remotely sensed measures and derived products; (4) with the addition of OSM road network features. The results show an improvement of the model fit with the addition of each of the predictive variables groups (Table 4). While GUF and WSF together explain 66% of the fit, the addition of nighttime lights improves the fit of the model (indicated by explanation of up to 76% of the variation). The addition

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

of further remotely sensed measures (i.e. Sentinel-2-derived spectral indices, slope, texture and forest cover) improves the model fit by a further 5% (up to 81% of the variation). Interestingly, the addition of OSM transportation network features only improves the fit of the model marginally to around 82%.

**Table 3**. Model fit for the stepwise regression analysis of nine evaluated predictors

| Step | Variable | $R^2$ | Adjusted $R^2$ | C(p) | AIC | RMSE |
|---|---|---|---|---|---|---|
| 1 | WSF | 0.614 | 0.613 | 984.9 | 18235.2 | 13332.1 |
| 2 | UI | 0.705 | 0.704 | 559.9 | 18013.3 | 11666.7 |
| 3 | GUF | 0.764 | 0.763 | 282.9 | 17828.1 | 10435.7 |
| 4 | VIIRS | 0.799 | 0.798 | 120.2 | 17696.0 | 9636.3 |
| 5 | Road length | 0.814 | 0.813 | 49.5 | 17631.2 | 9264.0 |
| 6 | FC area | 0.820 | 0.819 | 23.4 | 17605.8 | 9118.9 |
| 7 | NDBI | 0.822 | 0.821 | 17.0 | 17599.5 | 9078.8 |
| 8 | Number of junctions | 0.823 | 0.822 | 13.1 | 17595.6 | 9052.3 |
| 9 | Median slope | 0.614 | 0.613 | 11.9 | 17594.3 | 9040.2 |

**Table 4**. Four regression model outputs showing an improvement of model fit with the inclusion of four groups of variables: (1) only GUF and WSF; (2) addition of nighttime lights (VIIRS); (3) addition of remotely sensed measures and derived products; (4) addition of OSM road network features.

| Step | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| GUF | 0.115*** | 0.124*** | 0.127*** | 0.138*** |
| | (0.010) | (0.009) | (0.008) | (0.008) |
| WSF | 0.141*** | 0.081*** | 0.050*** | 0.034*** |
| | (0.010) | (0.009) | (0.009) | (0.009) |
| VIIRS | | 2,214.671*** | 1,276.821** | 1,073.838*** |
| | | (124.738) | (127.805) | (126.619) |
| NDBI | | | -37.152*** | -17.790 |
| | | | (11.258) | (11.409) |
| NDVI | | | 72,452.840** | 64,046.550** |
| | | | (30,894.100) | (29,957.530) |
| SAVI | | | -48,312.220** | -42,708.300** |
| | | | (20,600.640) | (19,976.140) |
| UI | | | 46.857*** | 25.415** |
| | | | (9.751) | (10.191) |
| Forest cover | | | 0.060*** | 0.051*** |
| | | | (0.012) | (0.011) |
| Slope | | | 179.453 | 274.377 |
| | | | (192.545) | (187.222) |
| Sentinel-1 | | | -696.229 | -295.342 |
| | | | (453.517) | (442.270) |
| Road length | | | | 1.470*** |
| | | | | (0.543) |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

| | | | | |
|---|---|---|---|---|
| No. of junctions | | | | 60.057** |
| | | | | (29.311) |
| Constant | 0.716 | -699.988 | 30,909.080*** | 17,403.480*** |
| | (672.122) | (573.990) | (3,897.596) | (4,351.436) |
| Observations | 835 | 835 | 835 | 835 |
| $R^2$ | 0.663 | 0.756 | 0.813 | 0.825 |
| Adjusted $R^2$ | 0.662 | 0.755 | 0.811 | 0.823 |
| Residual Std. Error | 12,460.39 | 10,615.940 | 9,329.420 | 9,029.806 |
| F Statistic | 818.245*** | 856.591*** | 357.937*** | 323.203*** |
| *Note:* | | | | *p<0.1; **p<0.05; ***p<0.01 |

*3.2. Prediction of OSM building footprint coverage*

The results above indicate that the area of OSM building footprints in a cell can be explained by several of the remotely sensed indicators. To evaluate the potential of these indicators to predict the area of OSM building footprints in a cell, we perform a regression with Random Forests. Random Forests [57] are tree-based models that include *k* decision trees and *p* randomly chosen predictors for each recursion. When predicting for an example, its variables are run through each of the *k* trees, and the *k* predictions are averaged through an arithmetic mean. The learning process of the forest involves some level of randomness. Each tree is trained using a subset of examples from the training set, drawn randomly with replacement, with each node's binary question determined using a random subset of *p* input variables. We perform the regression with the 835 cells that are visually assessed as relatively fully mapped (i.e. more than 75% of the buildings in a cell are assessed as mapped). To evaluate the accuracy of the prediction, we adopt a 5-fold cross validation method. In each experiment, the examples in one of the data folds are left out for testing and the examples in the remaining 4 folds are used to train the model. The performance quality of the trained model is tested on the examples in the left-out fold, and the overall performance measure is then averaged over the 5 folds. We assess classification accuracy with different number of decision trees: 2, 4, 8, 16, 32, 64, 128, 256 and 512, with minimum size of terminal nodes set to 5.

Random Forest regression predicts up to 89% of the variation of OSM building footprints in a cell. Performance improves with the addition of decision trees up to 64 trees, for example, from 81% to 89% of the predicted area (with 2 and 64 decision trees, respectively, Figure 5). Figure 6 presents a comparison between the actual and the predicted area of OSM building footprints in a cell (regression with 64 decision trees) (Figure 6a). The two most important variables to the model are WSF and GUF, followed by OSM road network features and Sentinel-2 derived spectral indices (indicated by variable importance sensitivity (*IncNodePurity*), see Figure 6b).

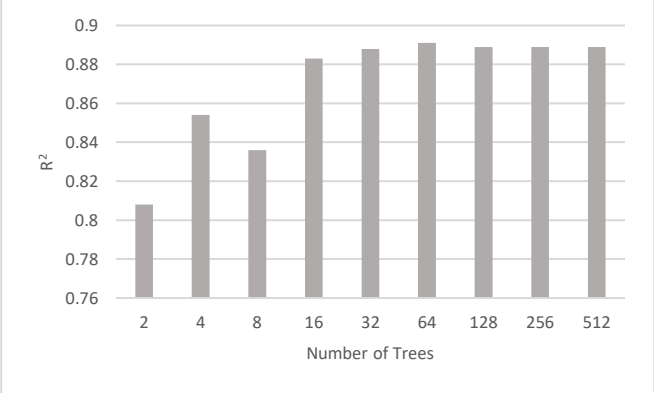
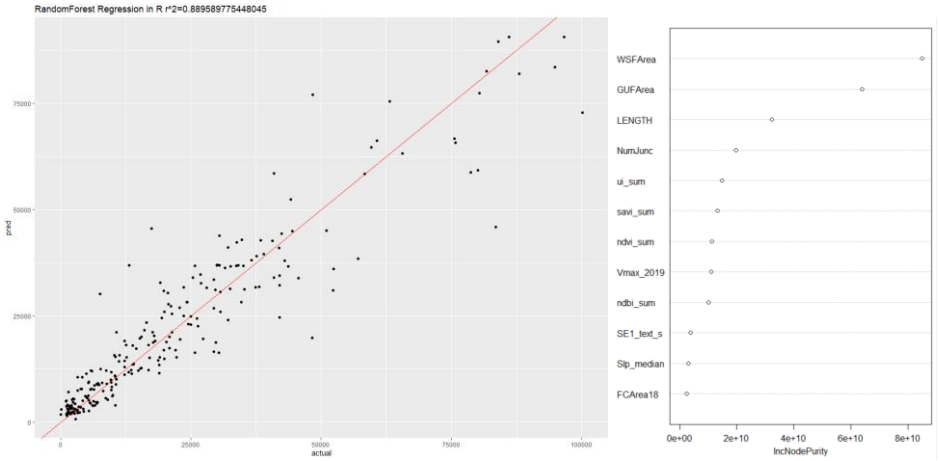



**Figure 5.** The improvement of $R^2$ with the increase from 2 to 64 decision trees in the Random Forest model




(a) (b)

**Figure 6.** Comparison between the (a) actual and predicted area of OSM building coverage in each cell (using 64 decision trees) and (b) variable importance sensitivity.

We use the Random Forest model to predict the area of OSM building footprints over the entire Haiti dataset (i.e. we train the model with the 835 cells assessed as relatively fully mapped and predict for the entire dataset). Figure 7b shows the predicted area of OSM building footprints per cell over Haiti. The results highlight large portions in Haiti that have not yet been mapped (e.g. for example, the northern cities Gonaives and Cap-Haitien) as well as patches of unmapped cells around major cities (e.g. Port-au-Prince). This analysis allows us to identify areas (cells) that are predicted to incorporate large areas of building footprints but actually lack coverage (Figure 8).

A visual examination shows that the predicted coverage of OSM building footprints (Figure 7b) corresponds more closely with the distribution of built-up land cover (according to GUF, for example) and nighttime lights (VIIRS) (Figures 7c and 7d, respectively) compared to the current distribution of OSM building footprints (Figure 7a).

To further evaluate the accuracy of the model, we perform an Ordinary Least Squares regression (OLS) analysis using the entire Haiti dataset (136,747 cells). We find that the remotely sensed indicators explain up to 89% of the variation of the predicted area of OSM building footprints in all the cells spanning Haiti ($R^2 = 0.89$, $F(12,136,734) = 90,690$, $p < 0.01$). In comparison, these indicators explain only 48% of the variation of the current area of OSM building footprints in the Haiti dataset ($R^2 = 0.48$, $F(12,136,734) = 10,330$, $p < 0.01$).

Finally, in order to identify cells that are predicted to incorporate building footprints but are not actually mapped, we calculate the ratio between the actual and the predicted area of OSM building footprint in a cell (calculated as the predicted area of OSM building footprints in a cell divided by the actual area of OSM building footprints in a cell) (Figure 9a). This analysis allows us to identify cells where the ratio between the predicted and the actual area of OSM building footprint in a cell is low, highlighting cells that require significant mapping (Figure 9b).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

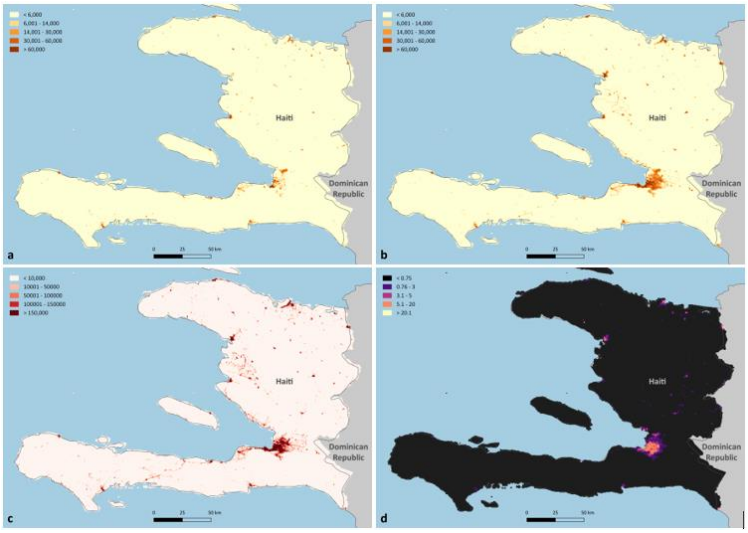



**Figure 7.** Actual (a) and predicted (b) area of OSM building footprints in a cell, compared to per-cell (c) area of GUF built-up land cover and (d) VIIRS SOL.

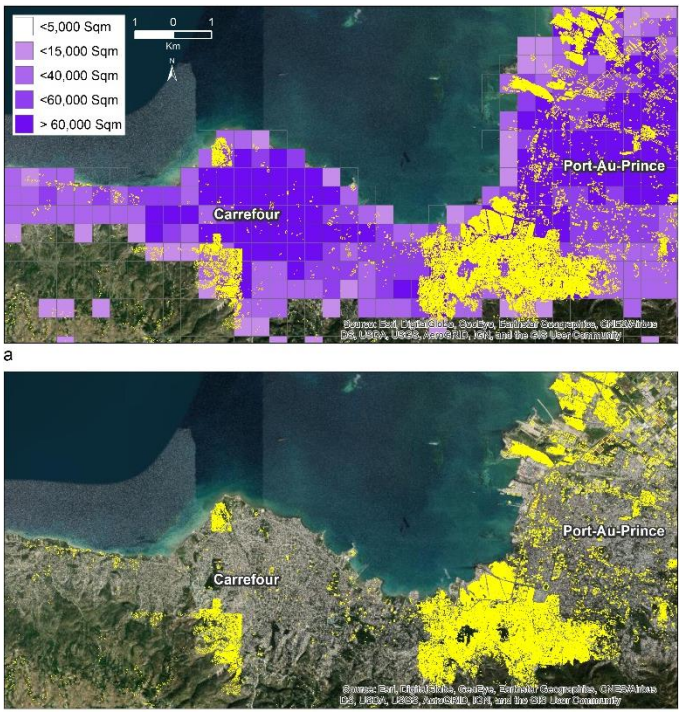



**Figure 8.** Predicted area of OSM building footprints in a cell (Carrefour). Areas shown in purple are predicted to include a large area of building footprints.

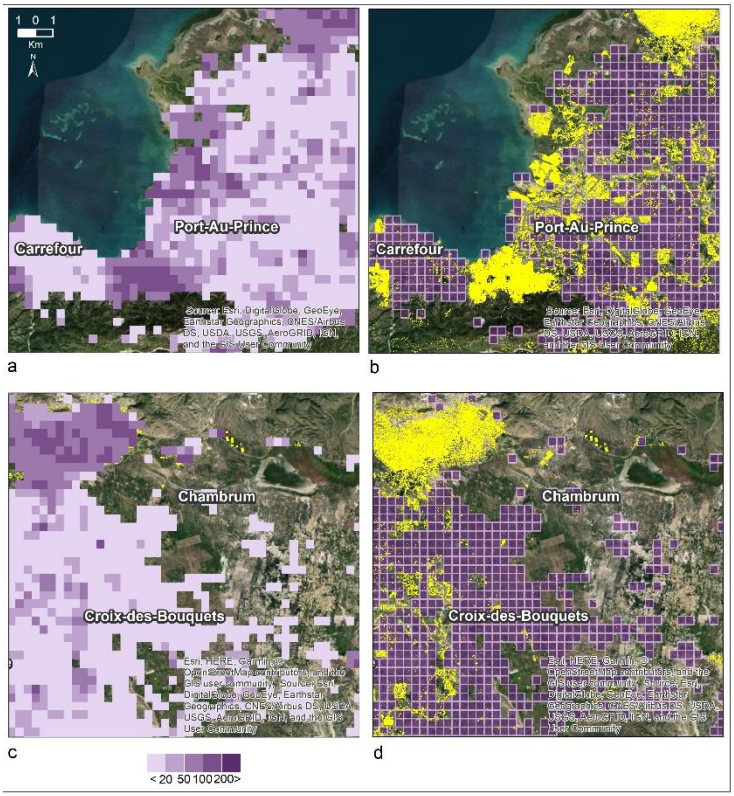



**Figure 9.** The ratio between the actual and the predicted area of OSM building footprint in a cell (calculated as the predicted area of OSM building footprints in a cell divided by the actual area of OSM building footprints in a cell) (a, c); Cells with the lowest ratio (lower than 20). Cells shown in purple are cells where the actual area of OSM building footprints is much lower than the predicted area.

### *3.3. Evaluation of the method in the case of Dominica and St. Lucia*

The results above suggest the potential of several remotely sensed indicators to predict the coverage of OSM building footprints, at least in the case of Haiti. In order to assess the validity of the method, we perform additional analysis in the case of two additional Small Island States: Dominica (area of 751 km$^2$) and St. Lucia (area of 616 km$^2$) (Figure 10). Dominica (approximately 74,000 people [59]) is located approximately 1,200 km southeast of Haiti, with large portions of its population residing in the capital Roseau (population 14,700) and Portsmouth (population 5,200) [59]. While smaller in size, St. Lucia has more than double the population (approximately 166,000 people [58]) of Dominica. A visual examination suggests that the coverage of OSM building footprints in Dominica is relatively complete, while large portions of St. Lucia remain unmapped (38,619 and 18,101 buildings have been mapped in Dominica and St. Lucia, respectively). Areas lacking OSM building footprints include parts of the capital, Castries, and the second largest town, Vieux Fort (See Figure 11).These two areas account for approximately 49% of the population of St. Lucia (64,654 and 16,624 people respectively [60]).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 November 2019 doi:10.20944/preprints201911.0361.v1

Peer-reviewed version available at Remote Sens. 2020, 12, 118; doi:10.3390/rs12010118

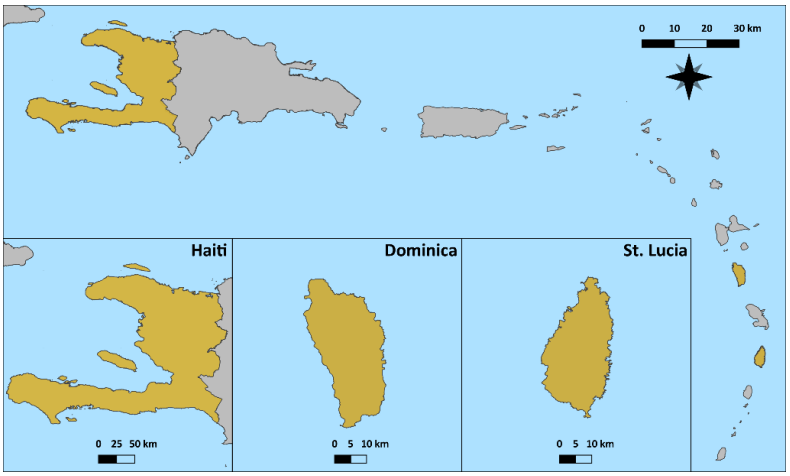



**Figure 10.** Locations and size comparisons of the three study areas: Haiti, Dominica, and St. Lucia

(a) Vieux Fort (b) Castries

**Figure 11.** Examples of areas missing OSM building footprints in St. Lucia (Vieux Fort (a) and the Castries (b))

Similar to the methodology described in the case of Haiti, we create a fishnet of cells, 0.25 sq. km in size, spanning the two islands (totaling 3,861 cells over Dominica and 2,796 cells over St. Lucia). In the case of Dominica, we find a high positive and significant correlation between the area of OSM building footprints in a cell and two predictors: GUF and VIIRS (r= 0.91 and r= 0.75, respectively, p<0.01 for both) and a lower (but significant) correlation with the four Sentinel-2 derived spectral indices: NDBI and UI (r=0.38, r=0.35, respectively, p<0.01) and NDVI and SAVI (r=-0.20, p<0.01 for both). An OLS regression analysis reveals that together, these variables explain 85% of the variation of OSM building footprint area in a cell ($R^2$= 0.85, $F(9,3851) = 2424$, $p < 0.01$). Random Forest regression (with 64 decision trees) results in similar trends, indicated by a high accuracy rate of around 85% (regression accuracy assessed using 5-fold cross validation). In the case of St. Lucia, the correlation between the area of OSM building footprints, GUF and WSF ranges between 0.70 and 0.75 (p<0.01). The correlation between OSM building footprints area and VIIRS is lower (r=0.58, p<0.01). Together, the predictors explain only 66% of the variation is OSM building footprints ($R^2$= 0.66, $F(12,2783) = 464.6$, $p < 0.01$). We relate the lower fit of the model to the fact that large portions of the country have not been mapped, resulting in significant "noise". Thus, we visually assess the completeness of OSM building footprints in St. Lucia cells, and identify 180 cells in which more than 75% of their area is assessed as mapped. With these visually assessed cells, the fit of the model improves, and together, the predictors explain 92% of the variation of OSM building footprint area ($R^2$=92%, $F(12,166) = 166.4$, $p < 0.01$). Random Forest regression (with 64 decision trees) results in a similar accuracy rate ($R^2$=92%) (Table 5).

**Table 5**. Pearson Correlation Test, OLS Regression, and Random Forest Regression for Dominica and St. Lucia

| | Dominica | St. Lucia | |
|---|---|---|---|
| | Full dataset (N=3861) | Full (N=2781) | Visually assessed cells* (N=179) |
| *(I) Pearson Correlation Test* | | | |
| GUF | r=0.91* | r=0.75* | r=0.89* |
| VIIRS | r=0.75* | r=0.58* | r=0.72* |
| NDBI | r=0.38* | r=0.30* | r=0.30* |
| UI | r=0.35* | r=0.26* | r=0.35* |
| NDVI | r=-0.20* | r=-0.19* | r=-0.29* |
| SAVI | r=-0.20* | r=-0.19* | r=-0.29* |
| *(II) OLS* | | | |
| | $R^2$=85% $F(9,3851) = 2424$, $p = 0.00$ | $R^2$=66% $F(12,2783) = 464.6$, $p = 0.00$ | $R^2$=92% $F(12,166) = 166.4$, $p = 0.00$ |
| *(III) Random Forest* | | | |
| | $R^2$=85% | | $R^2$=94% |

*Note: *p<0.01*

## 4. Discussion

In recent decades, natural disasters have been responsible for an estimated 0.1% of global deaths, killing on average 60,000 people per year [61]. In the last two decades, developing countries have accounted for more than half of all reported casualties [62]. Natural disasters often cause significant damage to communities, infrastructure and the environment, and require immediate intervention and implementation of appropriate measures aiming to save lives.

Accurate and easily accessible geospatial information is key for an effective disaster risk management cycle and for informed decision-making during humanitarian response [63]. The increasing availability of geospatial information is revolutionizing disaster research and emergency management. Until recently, much of this essential geospatial information was proprietary, scarce and in many cases, unavailable during significant disasters. Since the early 2000, however, Volunteered Geographic Information (VGI) has been playing a growing role in the support of humanitarian relief [63]. OSM, the world's first openly-licensed geospatial database created by volunteers, has repeatedly been shown to be highly suitable for disaster mapping [64] and has previously been used for disaster management in numerous major disasters, including the Nepal earthquake 2015 [18], the 2010 earthquake in Haiti [22] and 2016 Hurricane Matthew [23]. Recently, large corporations including Apple, Microsoft, and Facebook have been hiring editors to contribute to the OSM database [20].

Despite the continuous efforts to improve completeness of OSM database, large portions of the world remain unmapped, including in countries prone to natural hazards, reflecting limited internet speed connectivity, limited availability of GPS devices, lack of technical skilled volunteers and limited awareness to VGI technologies [19]. Often, when disasters occur, there is a lack of this essential information, which results in mapping campaigns, including Mapathons [65] that are designed to map the impacted areas. In the context of natural disasters, the coordination of volunteers' mapping efforts is operated by the Humanitarian OpenStreetMap Team (HOT), originally formed right after the Haiti earthquake. Other initiatives, such as Missing Maps were established in order to "map the most vulnerable places in the world", where mapping projects are

split into small tasks, allowing remote volunteers to work simultaneously on the same overall area (as of 2018, there were nearly Missing Maps 60,000 mappers) [66].

Several methods have been proposed to assess the completeness of OSM database, including evaluation of the completeness of street networks [67] [29], land use or buildings [68] [68]. Generally, there are pronounced differences in completeness between urban and rural areas, with the latter being less well-covered by OSM data [69]. Several tools have also been developed to visually assess the relative completeness of OSM mapping features. For example, OSMMatrix [70] provides a web-based application which portrays the spatial distribution of key attributes derived from the OSM dataset. Another tool for gridded estimation of the relative completeness of OSM was proposed by Development Seed, comparing OSM building data with WorldPop (population density distribution per 100x100m) [71]. Approaches for estimation of OSM completeness include comparison between OSM mapped features and other existing datasets, for example, national administrative data [72] [73] [74] [75]. Because reference data is often unavailable and varies by countries, there is a need for a scalable methodology that would allow to rapidly assess the completeness of OSM features without the need for administrative data. The increased availability of free and open source remotely sensed data can be utilized to identify locations of built-up land cover- areas that require mapping of built-up features. To the best of our knowledge, no study has yet evaluated the potential use of remotely sensed measurements to predict the completeness of OSM features. In this study, we propose a methodology that relies on remotely sensed measurements and classification schemes to predict the completeness of OSM building footprints in the case of three small island states (Haiti, Dominica and St. Lucia). We find that the combinatorial effects of the predictors explain up to 94% of the variation of OSM building footprints completeness. Importantly, we find that the addition of various remotely sensed measures to existing classification schemes (i.e. GUF and WSF) as predictors – including nighttime lights and Sentinel-2 derived spectral indices - improves the prediction of the area of OSM building footprints. As more and more remotely sensed data become available to the research community, our study extends the existing literature by demonstrating how they could be leveraged to conduct novel, and critical, large-scale assessments of the completeness of OSM database, especially in areas at high risk to natural disasters.

## 5. Conclusions

Globally, there has been an increase in the frequency and impacts of major natural disaster events and over the next century it is likely that climate change will amplify the number and severity of such disasters. While accurate and timely geospatial information is vital for the full cycle of disaster risk management, this data is not always available for the disaster management community when disasters occurs. Although VGI platforms, specifically OpenStreetMap (OSM), show great potential to support humanitarian mapping tasks, gaps in VGI data remains a major concern [76]. There is an increasing need for a fully automatic tool that would allow to identify areas that lack a complete mapping of OSM features– especially in areas prone to hazard events. W-hile previous studies have utilized OSM data as reference for classification of built-up land cover with satellite imagery [77] [78] [79] here we show the potential use of publicly available remotely sensed data as predictors of the spatial coverage of OSM building footprints.

With the increase in the frequency and severity of disaster events – especially in developing countries – it is essential to ensure the availability, accessibility and accuracy of geographical information for disaster management operations. The tool and methodology we present here are time efficient and scalable.

Extension to our approach may improve the accuracy of the prediction of OSM building footprints area by adding additional remotely sensed measures. Incorporating additional datasets such as newly developed VIIRS nighttime light products, socio-economic variables, additional land cover and land use classification schemes may offer opportunities to improve the accuracy of the prediction.

**Author Contributions:** Ran Goldblatt designed the experiments, together with Nick Jones and Jenny Mannix, the methodology, implemented the experiment and wrote the manuscript. Jenny Mannix and Nick Jones helped to carry out and implement the experiment and improve the manuscript.

**Funding:** This research was funded by the World Bank Group.

**Acknowledgments:** We thank Brad Bottoms for the excellent help throughout this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. WHO Available online: https://www.who.int/environmental_health_emergencies/natural_events/en/ (accessed on Nov 1, 2019).
2. FAO Available online: http://www.fao.org/resilience/areas-of-work/natural-hazards/en/ (accessed on Nov 1, 2019).
3. Munang, R.; Thiaw, I.; Alverson, K.; Liu, J.; Han, Z. The role of ecosystem services in climate change adaptation and disaster risk reduction. *Curr. Opin. Environ. Sustain.* **2013**, *5*, 47–52.
4. Brown, P.; Daigneault, A.J.; Tjernström, E.; Zou, W. Natural disasters, social protection, and risk perceptions. *World Dev.* **2018**, *104*, 310–325.
5. Shepherd, A.; Mitchell, T.; Lewis, K.; Lenhardt, A.; Jones, L.; Scott, L.; Muir-Wood, R. *The geography of poverty disasters and climate extremes in 2030*; 2013;
6. Shen, S.; Cheng, C.; Song, C.; Yang, J.; Yang, S.; Su, K.; Yuan, L.; Chen, X. Spatial distribution patterns of global natural disasters based on biclustering. *Nat. Hazards* **2018**, *92*, 1809–1820.
7. Eckstein, D.; Künzel, V.; Schäfer, L. Global climate risk index 2018. *Ger. Bonn* **2017**.
8. Fairbairn, T.I. Economic consequences of natural disasters among Pacific island countries. **2019**.
9. Van Beynen, P.; Akiwumi, F.A.; Van Beynen, K. A sustainability index for small island developing states. *Int. J. Sustain. Dev. World Ecol.* **2018**, *25*, 99–116.
10. Goodchild, M.F. *Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. IJSDIR 2: 24–32*; 2007;
11. Kawasaki, A.; Berman, M.L.; Guan, W. The growing role of web-based geospatial technology in disaster response and support. *Disasters* **2013**, *37*, 201–221.
12. Horita, F.E.A.; Degrossi, L.C.; de Assis, L.F.G.; Zipf, A.; de Albuquerque, J.P. The use of volunteered geographic information (VGI) and crowdsourcing in disaster management: a systematic literature review. **2013**.
13. Poorazizi, M.; Hunter, A.; Steiniger, S. A volunteered geographic information framework to enable bottom-up disaster management platforms. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1389–1422.
14. Chen, H.; Zhang, W.C.; Deng, C.; Nie, N.; Yi, L. Volunteered Geographic Information for Disaster Management with Application to Earthquake Disaster Databank & Sharing Platform. In Proceedings of the IOP Conference Series: Earth and Environmental Science; IOP Publishing, 2017; Vol. 57, p. 012015.
15. Mirbabaie, M.; Stieglitz, S.; Volkeri, S. Volunteered geographic information and its implications for disaster management. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS); IEEE, 2016; pp. 207–216.
16. Neis, P.; Zielstra, D. Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. *Future Internet* **2014**, *6*, 76–106.
17. Schelhorn, S.-J.; Herfort, B.; Leiner, R.; Zipf, A.; De Albuquerque, J.P. Identifying elements at risk from OpenStreetMap: The case of flooding. In Proceedings of the ISCRAM; 2014.

18. Poiani, T.H.; dos Santos Rocha, R.; Degrossi, L.C.; de Albuquerque, J.P. Potential of collaborative mapping for disaster relief: A case study of OpenStreetMap in the Nepal earthquake 2015. In Proceedings of the 2016 49th Hawaii International Conference on System Sciences (HICSS); IEEE, 2016; pp. 188–197.
19. Latif, S.; Islam, K.R.; Khan, M.M.I.; Ahmed, S.I. OpenStreetMap for the disaster management in Bangladesh. In Proceedings of the 2011 IEEE Conference on Open Systems; IEEE, 2011; pp. 429–433.
20. Palen, L.; Soden, R.; Anderson, T.J.; Barrenechea, M. Success & scale in a data-producing organization: The socio-technical evolution of OpenStreetMap in response to humanitarian events. In Proceedings of the Proceedings of the 33rd annual ACM conference on human factors in computing systems; ACM, 2015; pp. 4113–4122.
21. Eckle, M.; Herfort, B.; Yan, Y.; Kuo, C.-L.; Zipf, A. Towards using Volunteered Geographic Information to monitor post-disaster recovery in tourist destinations. In Proceedings of the ISCRAM; 2017.
22. Soden, R. 4 Years On, Looking Back at OpenStreetMap Response to the Haiti Earthquake. *Lat. Am. Caribb.* 2014.
23. Xu, J.; Li, L.; Zhou, Q. SPATIAL-TEMPORAL ANALYSIS OF OPENSTREETMAP DATA AFTER NATURAL DISASTERS: A CASE STUDY OF HAITI UNDER HURRICANE MATTHEW. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *42*.
24. Barrington-Leigh, C.; Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PloS One* **2017**, *12*, e0180698.
25. Hecht, R.; Kunze, C.; Hahmann, S. Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1066–1091.
26. Mashhadi, A.; Quattrone, G.; Capra, L. The Impact of Society on Volunteered Geographic Information: The Case of OpenStreetMap. In *OpenStreetMap in GIScience: Experiences, Research, and Applications*; Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M., Eds.; Lecture Notes in Geoinformation and Cartography; Springer International Publishing: Cham, 2015; pp. 125–141 ISBN 978-3-319-14280-7.
27. Quattrone, G.; Mashhadi, A.; Capra, L. Mind the Map: The Impact of Culture and Economic Affluence on Crowd-mapping Behaviours. In Proceedings of the Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing; ACM: New York, NY, USA, 2014; pp. 934–944.
28. Zielstra, D.; Zipf, A. A comparative study of proprietary geodata and volunteered geographic information for Germany. *13th AGILE Int. Conf. Geogr. Inf. Sci.* **2010**.
29. Haklay, M. How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. B* **2010**, *37*, 682–703.
30. Zielstra, D.; Hochmair, H.H.; Neis, P. Assessing the Effect of Data Imports on the Completeness of OpenStreetMap – A United States Case Study. *Trans. GIS* **2013**, *17*, 315–334.
31. Mocnik, F.-B.; Mobasheri, A.; Zipf, A. Open source data mining infrastructure for exploring and analysing OpenStreetMap. *Open Geospatial Data Softw. Stand.* **2018**, *3*, 7.

32. Antoniou, V.; Skopeliti, A. Measures and Indicators of Vgi Quality: AN Overview. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *3*, 345–351.
33. Acheson, E.; De Sabbata, S.; Purves, R.S. A quantitative analysis of global gazetteers: Patterns of coverage for common feature types. *Comput. Environ. Urban Syst.* **2017**, *64*, 309–320.
34. Zhang, Y.; Li, X.; Wang, A.; Bao, T.; Tian, S. Density and diversity of OpenStreetMap road networks in China. *J. Urban Manag.* **2015**, *4*, 135–146.
35. Törnros, T.; Dorn, H.; Hahmann, S.; Zipf, A. UNCERTAINTIES OF COMPLETENESS MEASURES IN OPENSTREETMAP – A CASE STUDY FOR BUILDINGS IN A MEDIUM-SIZED GERMAN CITY. In Proceedings of the ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences; Copernicus GmbH, 2015; Vol. II-3-W5, pp. 353–357.
36. Brovelli, M.; Zamboni, G. A new method for the assessment of spatial accuracy and completeness of OpenStreetMap building footprints. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 289.
37. Husen, S.N.R.M.; Idris, N.H.; Ishak, M.H.I. The Quality of Openstreetmap in Malaysia: a Preliminary Assessment. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *4249*, 291–298.
38. Brovelli, M.A.; Minghini, M.; Molinari, M.; Mooney, P. Towards an Automated Comparison of OpenStreetMap with Authoritative Road Datasets. *Trans. GIS* **2017**, *21*, 191–206.
39. Senaratne, H.; Mobasheri, A.; Ali, A.L.; Capineri, C.; Haklay, M. A review of volunteered geographic information quality assessment methods. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 139–167.
40. Luo, N.; Wan, T.; Hao, H.; Lu, Q. Fusing High-Spatial-Resolution Remotely Sensed Imagery and OpenStreetMap Data for Land Cover Classification Over Urban Areas. *Remote Sens.* **2019**, *11*, 88.
41. Audebert, N.; Saux, B.L.; Lefèvre, S. Joint Learning from Earth Observation and OpenStreetMap Data to Get Faster Better Semantic Maps. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2017; pp. 1552–1560.
42. World Bank The World Bank In Haiti Available online: https://www.worldbank.org/en/country/haiti/overview.
43. World Bank Understanding the Future of Haitian Cities Available online: https://projects.worldbank.org/en/results/2018/06/26/understanding-the-future-of-haitian-cities (accessed on Nov 5, 2019).
44. Bhawan, S.; Cohen, M. *Climate Change Resilience: The Case of Haiti*; OXFAM RESEARCH REPORTS; 2014;
45. Sam, J. Why is Haiti vulnerable to natural hazards and disasters? *theguardian* 2016.
46. UNFPA *Haiti : Humanitarian Action Fact Sheet*; Safety & Dignity for Women, Adolescent Girls & Young People, 2018;
47. NASA *NASA Visible Infrared Imaging adiometer Suite Level-1B Product User Guide [Collection-1]*; Level-1 and Atmosphere Archive & Distribution System; NASA Goddard Space Flight Center: Greenbelt, Maryland, 2018;

48. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.-M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510.
49. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309.
50. Zha, Y.; Gao, J.; Ni, S. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Int. J. Remote Sens.* **2003**, *24*, 583–594.
51. Kawamura, M. Relation between social and environmental conditions in Colombo Sri Lanka and the urban index estimated by satellite remote sensing data.; 1996; pp. 190–191.
52. Goldblatt, R.; Deininger, K.; Hanson, G. Utilizing publicly available satellite data for urban research: Mapping built-up land cover and land use in Ho Chi Minh City, Vietnam. *Dev. Eng.* **2018**, *3*, 83–99.
53. Esch, T.; Heldens, W.; Hirner, A.; Keil, M.; Marconcini, M.; Roth, A.; Zeidler, J.; Dech, S.; Strano, E. Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS J. Photogramm. Remote Sens.* **2017**, *134*, 30–42.
54. Esch, T.; Bachofer, F.; Heldens, W.; Hirner, A.; Marconcini, M.; Palacios-Lopez, D.; Roth, A.; Üreyen, S.; Zeidler, J.; Dech, S.; et al. Where We Live—A Summary of the Achievements and Planned Evolution of the Global Urban Footprint. *Remote Sens.* **2018**, *10*, 895.
55. Esch, T.; Schenk, A.; Ullmann, T.; Thiel, M.; Roth, A.; Dech, S. Characterization of Land Cover Types in TerraSAR-X Images by Combined Analysis of Speckle Statistics and Intensity Information. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1911–1925.
56. Republic of Haiti *POPULATION TOTALE, DE 18 ANS ET PLUS MENAGES ET DENSITES ESTIMES EN 2015*; Haitian Institute of Statistics and Information, Ministry of Finance, 2015;
57. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32.
58. Saint Lucia Open Data *2010 Census Population*; 2010;
59. Commonwealth of Dominica *2011 POPULATION AND HOUSING CENSUS: PRELIMINARY RESULTS*; CENTRAL STATISTICAL OFFICE, MINISTRY OF FINANCE, KENNEDY AVENUE, 2011;
60. Saint Lucia Open Data *2010 Census Population*; 2010;
61. Ritchie, H.; Roser, M. *Natural disasters*; Our World in Data; 2014;
62. Pelling, M.; Maskrey, A.; Ruiz, P.; Hall, P.; Peduzzi, P.; Dao, Q.-H.; Mouton, F.; Herold, C.; Kluser, S. Reducing disaster risk: a challenge for development. **2004**.
63. Giardino, M.; Perotti, L.; Lanfranco, M.; Perrone, G. GIS and geomatics for disaster management and emergency relief: a proactive response to natural hazards. *Appl. Geomat.* **2012**, *4*, 33–46.
64. Marco, M., Alessandro, S., Fabio, L., Napolitano, M., & Alessandro, P. Collaborative mapping response to disasters through OpenStreetMap: The case of the 2016 Italian earthquake. **2017**. The article is archived and publicly accessible with DOI on Zenodo at http://doi.org/10.5281/zenodo.1194529
65. Taylor, L.N. Digital Humanities Tools for Disaster Response: Hosting Mapathons and Telling Our Stories. **2018**.

66. Scholz, S.; Knight, P.; Eckle, M.; Marx, S.; Zipf, A. Volunteered Geographic Information for Disaster Risk Reduction—The Missing Maps Approach and Its Potential within the Red Cross and Red Crescent Movement. *Remote Sens.* **2018**, *10*, 1239.
67. Neis, P.; Zielstra, D.; Zipf, A. The street network evolution of crowdsourced maps: OpenStreetMap in Germany 2007–2011. *Future Internet* **2012**, *4*, 1–21.
68. Dorn, H.; Törnros, T.; Zipf, A. Quality Evaluation of VGI Using Authoritative Data—A Comparison with Land Use Data in Southern Germany. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 1657–1671.
69. Jacobs, C.; Zipf, A. Completeness of citizen science biodiversity data from a volunteered geographic information perspective. *Geo-Spat. Inf. Sci.* **2017**, *20*, 3–13.
70. Roick, O.; Hagenauer, J.; Zipf, A. OSMatrix–grid-based analysis and visualization of OpenStreetMap. *State Map EU* **2011**.
71. Anwar, S. Map Completeness and OSM Analytics. *Medium* 2018.
72. Hecht, R.; Kunze, C.; Hahmann, S. Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1066–1091.
73. Zheng, S.; Zheng, J. Assessing the completeness and positional accuracy of OpenStreetMap in China. In *Thematic cartography for the society*; Springer, 2014; pp. 171–189.
74. Kunze, C.; Hecht, R.; Hahmann, S. Assessing the completeness of building footprints in OpenStreetMap: an example from Germany.; 2013; pp. 25–30.
75. Antunes, F.; Fonte, C.C.; Brovelli, M.A.; Minghini, M.; Molinari, M.E.; Mooney, P. Assessing OSM road positional quality with authoritative data.; PRT, 2015; pp. 1–8.
76. Barron, C.; Neis, P.; Zipf, A. A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Trans. GIS* **2014**, *18*, 877–895.
77. Grippa, T.; Georganos, S.; Zarougui, S.; Bognounou, P.; Diboulo, E.; Forget, Y.; Lennert, M.; Vanhuysse, S.; Mboga, N.; Wolff, E. Mapping urban land use at street block level using openstreetmap, remote sensing data, and spatial metrics. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 246.
78. Brinkhoff, T. OPEN STREET MAP DATA AS SOURCE FOR BUILT-UP AND URBAN AREAS ON GLOBAL SCALE. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*.
79. Hu, T.; Yang, J.; Li, X.; Gong, P. Mapping urban land use by using landsat images and open social data. *Remote Sens.* **2016**, *8*, 151.

[i] https://wiki.openstreetmap.org/wiki/Stats